

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-127985

(43)Date of publication of application : 16.05.1997

(51)Int.Cl.

G10L 7/04

G10L 9/14

G10L 9/16

H03M 7/30

(21)Application number : 07-302128 (71)Applicant : SONY CORP

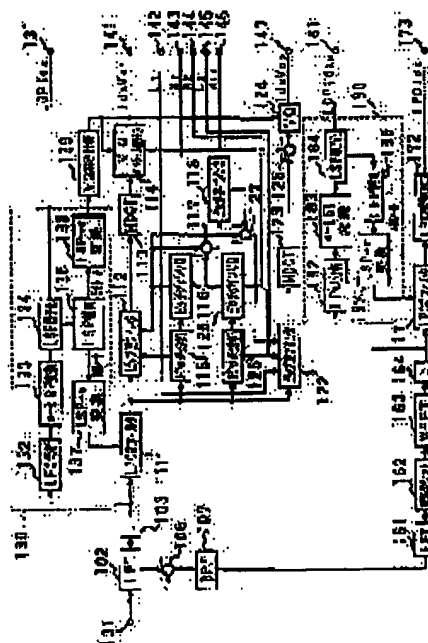
(22)Date of filing : 26.10.1995 (72)Inventor : OMORI SHIRO  
NISHIGUCHI MASAYUKI  
MATSUMOTO ATSUSHI  
IIJIMA KAZUYUKI

## (54) SIGNAL CODING METHOD AND DEVICE THEREFOR

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To perform a regeneration of high quality at a low bit rate by band-dividing an input signal and coding it by respective methods suitable to low-pass side and high-pass side

**SOLUTION:** From the input signal from a terminal 101, a high-pass side signal is taken out by a LPF (low pass filter) 102 and a subtracter 106, and this signal is FFT-processed by a FFT(fast Fourier-transformation) circuit 161, and shifted to a low-pass side band by a frequency shifting circuit. The signal reverse-FFT processed by a reverse FFT circuit 163 is transmitted to a LPC(linear predictive coding) reverse filter circuit 171 and subjected to predictive coding processing, whereby the predictive coding processing can be realized at low sampling frequency.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's  
decision of rejection]

[Kind of final disposal of application  
other than the examiner's decision of  
rejection or application converted  
registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's  
decision of rejection]

[Date of requesting appeal against  
examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

(11)特許出願公開番号

特開平9-127985

(43)公開日 平成9年(1997)5月16日

(51)Int.Cl. <sup>6</sup>	識別記号	序内整理番号	F I	技術表示箇所
G 1 0 L	7/04		G 1 0 L	G
	9/14			J
	9/16			
H 0 3 M	7/30	9382-5K	H 0 3 M	A

審査請求 未請求 請求項の数6 FD (全 23 頁)

(21)出願番号	特願平7-302128	(71)出願人	000002185 ソニー株式会社 東京都品川区北品川6丁目7番35号
(22)出願日	平成7年(1995)10月26日	(72)発明者	大森 士郎 東京都品川区北品川6丁目7番35号 ソニ ー株式会社内
		(72)発明者	西口 正之 東京都品川区北品川6丁目7番35号 ソニ ー株式会社内
		(72)発明者	松本 淳 東京都品川区北品川6丁目7番35号 ソニ ー株式会社内
		(74)代理人	弁理士 小池 晃 (外2名)

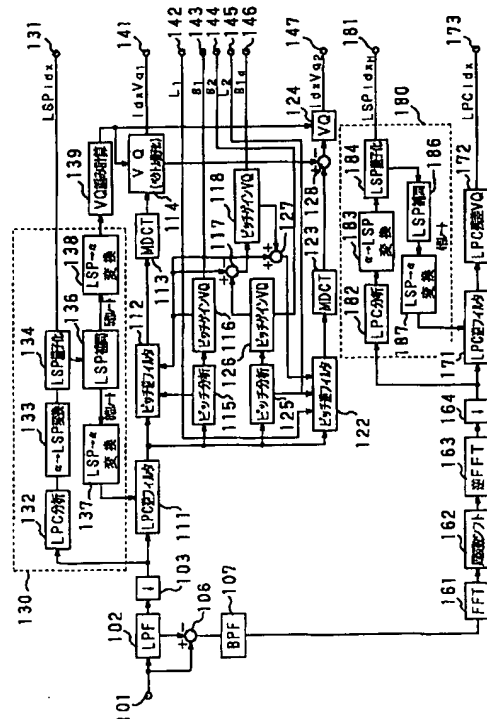
最終頁に続く

(54)【発明の名称】 信号符号化方法及び装置

(57) 【要約】

【課題】 入力信号を帯域分割し、低域側と高域側とでそれぞれ適した方法で符号化し、低ビットレートで高品質な再生を可能にする信号符号化方法及び信号符号化装置を提供する。

【解決手段】 端子１０１からの入力信号から、ＬＰＦ（ローパスフィルタ）１０２及び減算器１０６により高域側信号を取り出し、この信号をＦＦＴ（高速フーリエ変換）回路１６１によりＦＦＴ処理し、周波数シフト回路１６２により低域側の帯域にシフトさせ、逆ＦＦＴ回路１６３により逆ＦＦＴ処理した信号をＬＰＣ（線形予測符号化）逆フィルタ回路１７１に送って予測符号化処理を施すことにより、低いサンプリング周波数で予測符号化処理を実現できる。



## 【特許請求の範囲】

【請求項1】 入力信号を帯域分割する帯域分割工程と、  
分割された少なくとも1つの高域側の信号を低域側に周波数変換する周波数変換工程と、  
低域側に変換された信号のサンプリングレートを低下させるサンプリングレート低下工程と、  
サンプリングレートが低下させられた信号を予測符号化する工程とを有することを特徴とする信号符号化方法。

【請求項2】 上記帯域分割工程は、広帯域の入力信号を、電話帯域と、この電話帯域よりも高域側とに分割することを特徴とする請求項1記載の信号符号化方法。

【請求項3】 上記周波数変換工程は、上記分割された少なくとも1つの高域側の信号に対して直交変換を行った後、低域側に周波数シフト処理を行い、さらに、逆直交変換を行うことを特徴とする請求項1記載の信号符号化方法。

【請求項4】 上記直交変換は、高速フーリエ変換であることを特徴とする請求項3記載の信号符号化方法。

【請求項5】 上記予測符号化する工程は、上記サンプリングレート低下させられた信号を、線形予測しこのとき得られる線形予測係数を表現するパラメータ及び予測残差を量子化したものを取り出すことを特徴とする請求項1記載の信号符号化方法。

【請求項6】 入力信号を帯域分割する帯域分割手段と、  
分割された少なくとも1つの高域側の信号を低域側に周波数変換する周波数変換手段と、  
低域側に変換された信号のサンプリングレートを低下させるサンプリングレート低下手段と、  
サンプリングレートが低下させられた信号を予測符号化する手段とを有することを特徴とする信号符号化装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、広帯域音声信号のような入力信号を帯域分割して符号化する信号符号化方法及び装置に関する。

## 【0002】

【従来の技術】 オーディオ信号（音声信号や音響信号を含む）の時間領域や周波数領域における統計的性質と人間の聴感上の特性を利用して信号圧縮を行うような符号化方法が種々知られている。この符号化方法としては、大別して時間領域での符号化、周波数領域での符号化、分析合成符号化等が挙げられる。

【0003】 音声信号等の高効率符号化の例として、ハーモニック（Harmonic）符号化、MBE（Multiband Excitation: マルチバンド励起）符号化等のサイン波分析符号化や、SBC（Sub-band Coding: 帯域分割符号化）、LPC（Linear Predictive Coding: 線形予測符号化）、あるいはDCT（離散コサイン変換）、MDC

T（モデファイドDCT）、FFT（高速フーリエ変換）等が知られている。

## 【0004】

【発明が解決しようとする課題】 ところで、従来において、広帯域信号を符号化する場合、帯域分割を行いサブバンド符号化を行う方法が存在していた。特に、低域側と高域側とに2分割する場合、サンプリング周波数 $f_s/4$ で分割を行う方法が一般的である。この $f_s/4$ の周波数で分割した後、高域側についてはダウンサンプリングが行われる。こうすることで、高域周波数成分がそのまま低域側にエイリアシングとして折り返されるので、低域側の変換された信号として扱うことができたのである。

【0005】 ところが、 $f_s/4$ の周波数で分割できない場合においては、この方法をとることができなかった。低域側の信号を $f_s/2$ で再生することをも考えると、低域側は $f_s/4$ の周波数までは伸びていない。従って、高域側でこの部分をもカバーしなくてはならない。

【0006】 また、近年において、そのような帯域で分割される音声／楽音信号を符号化して、低域側信号を符号化したものと重ねてスケラビリティを持つビットストリームを作成する方法が望まれている。

【0007】 そこで、本発明は、このような実情に鑑みてなされたものであり、入力信号を帯域分割し、低域側と高域側とでそれぞれ適した方法で符号化することを可能にする信号符号化方法及び信号符号化装置を提供することを目的とする。

## 【0008】

【課題を解決するための手段】 本発明に係る信号符号化方法は、上述した課題を解決するために、入力信号を帯域分割し、分割された少なくとも1つの高域側の信号を低域側に周波数変換し、低域側に変換された信号のサンプリングレートを低下させ、サンプリングレートが低下させられた信号を予測符号化することを特徴としている。

【0009】 また、本発明に係る信号符号化装置は、上述した課題を解決するために、入力信号を帯域分割する帯域分割手段と、分割された少なくとも1つの高域側の信号を低域側に周波数変換する周波数変換手段と、低域側に変換された信号のサンプリングレートを低下させるサンプリングレート低下手段と、サンプリングレートが低下させられた信号を予測符号化する手段とを有することを特徴としている。

## 【0010】

【発明の実施の形態】 以下、本発明に係る好ましい実施の形態について説明する。まず、図1は、本発明に係る信号符号化方法の実施の形態が適用された広帯域音声信号の符号化装置を示している。

【0011】 ここで、図1の符号化装置の基本的な考え方は、入力信号を複数の帯域に分割して、分割された帯

10

20

30

40

50

## 3

域毎の信号に対してそれぞれの帯域の信号特性に応じた互いに異なる符号化を施すものである。具体的には、広帯域の入力音声信号を複数の帯域、例えば音声として充分な明瞭度を得られる電話帯域と、この電話帯域よりも高域側の帯域とに分割している。低域側の電話帯域の信号については、LPC（線形予測符号化）分析等の短期予測の後にピッチ予測等の長期予測を行った上で直交変換を施し、この変換後の係数を聴覚重み付けベクトル量子化しており、また、LPC係数等の短期予測係数を表現するパラメータ、ピッチやピッチゲイン等の長期予測に関連した情報についても量子化している。電話帯域よりも高域の信号については、短期予測したものをそのまま時間軸上でベクトル量子化している。

【0012】上記直交変換としてMDCT（モディファイド離散コサイン変換）を用い、このときの変換長を短くすることでベクトル量子化の重み付けを容易に行い、さらにこの変換長を $2^N$ 、すなわち2のべき乗の大きさにすることでFFT（高速フーリエ変換）を用いた高速化を可能としている。また、上記短期予測の残差を算出し、直交変換係数をベクトル量子化する際の重み算出するためのLPC係数は（ポストフィルタについても）、現フレーム内で求められたものと過去のフレームで求められたものとを滑らかに補間したものを用い、各分析サブフレームで最適なLPC係数を用いている。また、上記長期予測を行う場合、1フレームに複数回の予測又は補間を行い、それらのピッチラグ、ピッチゲインをそのまま、あるいは差分をとった上で量子化し、あるいは補間方法を示すフラグを伝送している。さらに、予測回数（頻度）の増加に対して、分散の小くなる予測残差について、それらの直交変換係数の差を量子化するマルチステージ（多段）ベクトル量子化を行うか、あるいは分割された帯域の内、ある1つの帯域に対するパラメータのみを用い、1つのエンコードビットストリームの全体あるいは一部分により、レート異なる複数のデコード動作を可能としている。

【0013】以下、図1を参照しながら説明する。図1の入力端子101には、例えばサンプリング周波数 $F_s$ が16kHzで0~8kHz程度の広帯域音声信号が供給されている。この入力端子101からの広帯域音声信号は、ローパスフィルタ102及び減算器106により低域側、例えば0~3.8kHz程度のいわゆる電話帯域の信号と、高域側、例えば3.8kHz~8kHzの信号とに帯域分割される。低域側は、サンプリング周波数変換器103により、サンプリング定理を満たす範囲で間引くことにより、例えば8kHzサンプリングの信号にしておく。

【0014】低域側での処理としては、まず、LPC分析・量子化部130により、例えば1ブロック256サンプル程度の分析長によりハミング窓かけをした上で、10次程度のLPC係数、すなわち $\alpha$ パラメータを算出

## 4

し、LPC逆フィルタ111によりLPC残差を求めている。このLPC分析の際には、分析の単位となる1ブロック256サンプルの内の96サンプルを次のブロックとオーバーラップさせることにより、フレーム間隔すなわちフレームインターバルは160サンプルとなる。このフレーム間隔は、8kHzサンプリングで20msecになる。また、このLPC分析・量子化部130では、LPC係数である $\alpha$ パラメータをLSP（線スペクトル対）パラメータに変換して量子化したものを伝送するようにしている。

【0015】すなわち、LPC分析・量子化部130において、サンプリング周波数変換器103からの低域側信号が入力されるLPC分析回路132は、入力信号波形の256サンプル程度の長さを1ブロックとしてハミング窓をかけて、自己相関法により線形予測係数、いわゆる $\alpha$ パラメータを求めている。データ出力の単位となるフレーミングの間隔は、例えば160サンプルで20msecである。

【0016】LPC分析回路132からの $\alpha$ パラメータは、 $\alpha \rightarrow$ LSP変換回路133に送られて、線スペクトル対（LSP）パラメータに変換される。これは、直接型のフィルタ係数として求めた $\alpha$ パラメータを、例えば10個、すなわち5対のLSPパラメータに変換する。変換は例えばニュートン・ラプソン法等を用いて行う。このLSPパラメータに変換するのは、 $\alpha$ パラメータよりも補間特性に優れているからである。

【0017】 $\alpha \rightarrow$ LSP変換回路133からのLSPパラメータは、LSP量子化器134によりベクトル量子化あるいはマトリクス量子化される。このとき、フレーム間差分をとってからベクトル量子化、あるいは、複数フレーム分をまとめてマトリクス量子化してもよい。ここでは、20msecを1フレームとし、20msec毎に算出されるLSPパラメータを2フレーム分まとめてマトリクス量子化している。

【0018】このLSP量子化器134からの量子化出力、すなわちLSPベクトル量子化のインデクスは、端子131を介して取り出され、また量子化済みのLSPベクトルあるいは逆量子化出力は、LSP補間回路136に送られる。

【0019】LSP補間回路136は、LSP量子化器134で上記20msec毎にベクトル量子化されたLSPのベクトルの前フレームと現フレームとの組を補間し、後の処理で必要となるレートにするためのものであり、この例では、8倍のレートと5倍のレートにしている。8倍レートでは、2.5msec毎にLSPベクトルが更新されるようにする。これは、残差波形を分析合成処理すると、その合成波形のエンベロープは非常になだらかでスムーズな波形になるため、LPC係数が20msec毎に急激に変化すると異音を発生することがあるからである。すなわち、2.5msec毎にLPC係数が徐

## 5

々に変化してゆくようにすれば、このような異音の発生を防ぐことができる。

【0020】このような補間が行われた2.5 msec 毎のLSPベクトルを用いて入力音声の逆フィルタリングを実行するために、LSP→ $\alpha$ 変換回路137により、LSPパラメータを例えば10次程度の直接型フィルタの係数である $\alpha$ パラメータに変換する。このLSP→ $\alpha$ 変換回路137からの出力は、上記LPC残差を求めるためのLPC逆フィルタ回路111に送られ、このLPC逆フィルタ111では、2.5 msec 毎に更新される $\alpha$ パラメータにより逆フィルタリング処理を行って、滑らかな出力を得るようにしている。

【0021】また、LSP補間回路136で5倍レートで補間された4 msec 毎のLSP係数は、LSP→ $\alpha$ 変換回路138に送られて $\alpha$ パラメータに変換され、MDCT係数の量子化に使用する重み計算のためのVQ（ベクトル量子化）重み計算回路139に送られる。

【0022】LPC逆フィルタ111からの出力は、長期予測であるピッチ予測のためのピッチ逆フィルタ112及び122に送られる。

【0023】次に、長期予測について説明する。長期予測は、ピッチ分析により求められたピッチ周期あるいはピッチラグ分だけ時間軸上でずらした波形を元の波形から減算してピッチ予測残差を求めることにより行っており、この例では3点ピッチ予測によって行っている。なお、ピッチラグとは、サンプリングされた時間軸データのピッチ周期に対応するサンプル数のことである。

【0024】すなわち、ピッチ分析回路115では1フレームに1回の割合、すなわち分析長が1フレームでピッチ分析が行われ、ピッチ分析結果の内のピッチラグ $L_1$ はピッチ逆フィルタ112及び出力端子142に送られ、ピッチゲインはピッチゲインVQ（ベクトル量子化）回路116に送られる。ピッチゲインVQ回路116では、上記3点予測に対応する3点でのピッチゲインがベクトル量子化され、コードブックインデクス $g_1$ が出力端子143より取り出され、代表値ベクトルあるいは逆量子化出力がピッチ逆フィルタ115、減算器117、加算器127にそれぞれ送られる。ピッチ逆フィルタ112は、上記ピッチ分析結果に基づいて3点ピッチ予測されたピッチ予測残差を出力する。このピッチ予測残差は、直交変換手段である例えばMDCT回路113に送られ、MDCT処理された後、VQ（ベクトル量子化）回路114により聴覚重み付けベクトル量子化される。このVQ回路114では、VQ重み計算回路139からの出力により聴覚重み付けされたベクトル量子化が施され、その出力であるインデクス $IdxVq_1$ は、出力端子141より取り出される。

【0025】またこの例においては、予測精度を高めるために、さらに別系統のピッチ予測系となるピッチ逆フィルタ122、ピッチ分析回路124及びピッチゲイン

## 6

VQ回路126を設けている。すなわち、上記各ピッチ分析中心の中間位置にも分析中心を置くようにして、ピッチ分析回路125により1/2フレーム周期で分析を行うようにしている。ピッチ分析回路125からのピッチラグ $L_2$ はピッチ逆フィルタ122及び出力端子145に、ピッチゲインはピッチゲインVQ（ベクトル量子化）回路126にそれぞれ送られる。ピッチゲインVQ回路126では、3点のピッチゲインベクトルをベクトル量子化して量子化出力であるピッチゲインのインデクス $g_2$ を出力端子144に送り、その代表ベクトルあるいは逆量子化出力を減算器117に送っている。ここで、元のフレーム周期の分析中心の位置でのピッチゲインは、ピッチゲインVQ回路116からのピッチゲインに近い値と考えられるから、この位置でのピッチゲインについては、ピッチゲインVQ回路116、126からの各逆量子化出力の差を減算器117でとって、これをピッチゲインVQ回路118でベクトル量子化して得られるピッチゲイン差分のインデクス $g_{ld}$ を出力端子146に送っている。このピッチゲイン差分の代表ベクトルあるいは逆量子化出力を、加算器127に送り、ピッチゲインVQ回路126からの代表ベクトルあるいは逆量子化出力と加算したものをピッチゲインとしてピッチ逆フィルタ122に送っている。なお、出力端子143から得られるピッチゲインのインデクス $g_2$ は、上記中間位置でのピッチゲインのインデクスである。ピッチ逆フィルタ122からのピッチ予測残差は、MDCT回路123でMDCT処理され、これが減算器128に送られて、VQ（ベクトル量子化）回路114からの代表ベクトルあるいは逆量子化出力が減算され、その差分がVQ回路124に送られてベクトル量子化され、インデクス $IdxVq_2$ が出力端子147に送られる。このVQ回路は、VQ重み計算回路139からの出力により聴覚重み付けのベクトル量子化を施す。

【0026】次に、高域側の信号処理について説明する。

【0027】この高域側の信号処理は、基本的に、入力信号を帯域分割し、分割された少なくとも1つの高域側の信号を低域側に周波数変換し、低域側に変換された信号のサンプリングレートを低下させて、サンプリングレートが低下させられた信号を予測符号化するようにしている。

【0028】図1の入力端子101に供給された広帯域信号が減算器106に入力され、この広帯域信号から、LPF（ローパスフィルタ）102により取り出された低域側の信号、例えば0～3.8 kHz程度のいわゆる電話帯域の信号、が差し引かれる。これによって、減算器106からは、高域側の信号、例えば3.8 kHz～8 kHzの信号が出力される。ただし、現実のLPF102の特性等により、減算器106からの出力には、3.8 kHz以下の成分もわずかながら残っており、高域側の信号

処理は、3.5 kHz以上、あるいは3.4 kHz以上の成分に対して行うようにしている。

【0029】この高域側の信号は、減算器106からの例えば3.5 kHz～8 kHzの4.5 kHzの周波数幅を持つが、ダウンサンプリング等により周波数を低域側にシフトあるいは変換して信号処理を行うため、例えば4 kHz幅まで狭めることが必要とされる。ここで、後の低域側との合成を考慮し、3.5 kHz～4 kHz付近は聴感上敏感であるので、ここをカットせずに、音声信号の性質として成分あるいはパワーも少なく聴感上影響の少ない7.5 kHz～8 kHzの0.5 kHz分をLPFあるいはBPF（バンドパスフィルタ）107によりカットする。

【0030】次に、低域側への周波数変換を行うが、この例では、直交変換手段、例えばFFT（高速フーリエ変換）回路161を用いて周波数軸上のデータに変換し、この周波数軸上のデータを周波数シフト回路162によりシフトした後、逆直交変換手段である逆FFT回路164により逆FFT処理することにより実現している。

【0031】逆FFT回路164からは、入力信号の高域側の例えば3.5 kHz～7.5 kHzの信号が、0～4 kHzの低域側に変換された信号が取り出される。この信号はサンプリング周波数が8 kHzで表現できるので、ダウンサンプリング回路164によりダウンサンプリングしてサンプリング周波数8 kHzの3.5 kHz～7.5 kHzの帯域の信号とする。このダウンサンプリング回路164からの出力は、LPC逆フィルタ171及びLPC分析・量子化部180のLPC分析回路182にそれぞれ送られる。

【0032】LPC分析・量子化部180は、上記低域側のLPC分析・量子化部130とほぼ同様な構成を有しているため、簡単に説明する。

【0033】すなわち、LPC分析・量子化部180において、ダウンサンプリング回路164からの低域変換された信号が入力されるLPC分析回路182は、入力信号波形の256サンプル程度の長さを1ブロックとしてハミング窓をかけて、例えば自己相関法により線形予測係数、いわゆる $\alpha$ パラメータを求めている。LPC分析回路182からの $\alpha$ パラメータは、 $\alpha \rightarrow$  LSP変換回路183に送られて、線スペクトル対（LSP）パラメータに変換される。 $\alpha \rightarrow$  LSP変換回路183からのLSPパラメータは、LSP量子化器184によりベクトル量子化あるいはマトリクス量子化される。このとき、フレーム間差分をとってからベクトル量子化してもよい。あるいは、複数フレーム分をまとめてマトリクス量子化してもよい。ここでは、20 msec を1フレームとし、20 msec 毎に算出されるLSPパラメータをベクトル量子化している。

【0034】このLSP量子化器184からの量子化出力、すなわち高域側信号のLSPベクトル量子化のイン

デクスLSPidx<sub>H</sub>は、端子181を介して取り出され、また量子化済みのLSPベクトルあるいは逆量子化出力は、LSP補間回路186に送られる。

【0035】LSP補間回路186は、LSP量子化器184で上記20 msec 毎にベクトル量子化されたLSPのベクトルの前フレームと現フレームとの組を補間し、後の処理で必要となるレートにするためのものであり、この例では、4倍のレートにしている。

【0036】このような補間が行われた5 msec 毎のLSPベクトルを用いて入力音声の逆フィルタリングを実行するために、LSP $\rightarrow$  $\alpha$ 変換回路187により、LSPパラメータをLPC合成フィルタの係数である $\alpha$ パラメータに変換する。このLSP $\rightarrow$  $\alpha$ 変換回路187からの出力は、上記LPC残差を求めるためのLPC逆フィルタ回路171に送られ、このLPC逆フィルタ171では、5 msec 毎に更新される $\alpha$ パラメータにより逆フィルタリング処理を行って、滑らかな出力を得るようにしている。

【0037】LPC逆フィルタ171からのLPC予測残差出力は、LPC残差VQ（ベクトル量子化）回路172に送られてベクトル量子化され、その出力であるLPC残差のインデクスLPCidxが出力端子173より取り出される。

【0038】以上のような構成の信号符号化装置において、低域側の一部構成を独立したコーデックの符号化装置としたり、あるいは出力されるビットストリームの全体と一部とを切り換えることにより、ビットレートの異なる信号伝送やデコードを可能としている。

【0039】すなわち、図1の構成の各出力端子からの全てのデータを伝送するとき、伝送ビットレートは16 kbps（kビット/秒）となり、一部端子からのデータを伝送することにより6 kbpsの伝送ビットレートとなる。

【0040】あるいは、図1の全ての端子からの16 kbpsの全データを伝送、すなわち送信あるいは記録し、受信あるいは再生側で16 kbpsの全データをデコードすることにより、16 kbpsの高品質の音声信号が得られ、6 kbpsのデータをデコードすることにより、簡単なデコーダで6 kbpsに応じた品質の音声信号が得られる。

【0041】ここで、図1の構成においては、出力端子131、141～143からの出力データが6 kbpsのデータに相当し、さらに出力端子144～147、173、181からの出力データを加えることで、16 kbpsの全データが得られる。

【0042】次に、上記図1の信号符号化装置に対応する信号復号化装置について、図2を参照しながら説明する。

【0043】この図2において、入力端子200には、上記図1の出力端子131からの出力に相当するLSP

10

20

30

40

50



のベクトル量子化出力、いわゆるコードブックのインデクスLSPidxが供給されている。

【0044】このLSPのインデクスLSPidxは、LPCパラメータ再生部240のLSPの逆VQ（逆ベクトル量子化）回路241に送られてLSP（線スペクトル対）データに逆ベクトル量子化あるいは逆マトリクス量子化され、LSP補間回路242に送られてLSPの補間処理が施された後、LSP→ $\alpha$ 変換回路243でLPC（線形予測符号）係数である $\alpha$ パラメータに変換され、この $\alpha$ パラメータがLPC合成フィルタ215、225及びピッチスペクトラルポストフィルタ216、226に送られる。

【0045】また、図4の入力端子201、202、203には、上記図1の各出力端子141、142、143からのMDCT係数のベクトル量子化のインデクスIsxVq1、ピッチラグL1、ピッチゲインg1がそれぞれ供給されている。

【0046】入力端子201からのMDCT係数のベクトル量子化のインデクスIsxVq1は、逆VQ（逆ベクトル量子化）回路211に供給されて逆ベクトル量子化され、逆MDCT回路212により逆MDCT処理された後、重畳加算（オーバーラップアッド）回路213で重畳加算され、ピッチ合成フィルタ214に送られる。ピッチ合成回路214には、各入力端子202、203からのピッチラグL1、ピッチゲインg1が供給されている。このピッチ合成回路214で、上記図1のピッチ逆フィルタ112でのピッチ予測符号化の逆処理が施された後、LPC合成フィルタ215に送られ、LPC合成処理が施される。このLPC合成された出力は、ピッチスペクトラルポストフィルタ216に送られて、ポストフィルタ処理が施され、出力端子219より6 kbpsのビットレートに対応する音声信号として取り出される。

【0047】図4の入力端子204、205、206及び207には、上記図1の各出力端子144、145、146及び147からのMDCT係数のベクトル量子化のピッチゲインg2、ピッチラグL2、インデクスIsxVq2及びピッチゲインg1dがそれぞれ供給されている。

【0048】入力端子207からのMDCT係数のベクトル量子化のインデクスIsxVq2は、逆VQ回路220に供給されて逆ベクトル量子化され、加算器221に送られて逆VQ回路211からの逆ベクトル量子化されたMDCT係数と加算され、逆MDCT回路222により逆MDCT処理された後、重畳加算（オーバーラップアッド）回路223で重畳加算され、ピッチ合成フィルタ214に送られる。このピッチ合成フィルタ214には、各入力端子202、204、205からのピッチラグL1、ピッチゲインg2、ピッチラグL2が供給されると共に、入力端子203からのピッチゲインg1と入力端子206からのピッチゲインg1dとが加算器217で加算されたものが供給されている。このピッチ合成フィル

タ224でピッチ残差の合成処理が施された後、LPC合成フィルタ225に送られ、LPC合成処理が施される。このLPC合成された出力は、ピッチスペクトラルポストフィルタ226に送られて、ポストフィルタ処理が施され、アップサンプリング回路227に送られてサンプリング周波数が例えば8 kHzから16 kHzにアップサンプリングされた後、加算器228に送られる。

【0049】さらに、入力端子207には、図1の出力端子181からの高域側のLSPインデクスLSPidxHが供給されており、このLSPのインデクスLSPidxHは、LPCパラメータ再生部245のLSPの逆VQ（逆ベクトル量子化）回路246に送られてLSPデータに逆ベクトル量子化され、LSP補間回路247に送られてLSPの補間処理が施された後、LSP→ $\alpha$ 変換回路248でLPC係数の $\alpha$ パラメータに変換され、この $\alpha$ パラメータが高域側LPC合成フィルタ232に送られる。

【0050】入力端子209には、図1の出力端子173からの高域側のLPC残差のベクトル量子化出力であるインデクスLPCidxが供給されて、高域逆VQ回路231で逆ベクトル量子化され、高域側LPC合成フィルタ232に送られる。高域側LPC合成フィルタ232でLPC合成処理された出力は、アップサンプリング回路233でサンプリング周波数が例えば8 kHzから16 kHzにアップサンプリングされた後、直交変換手段であるFFT回路234で高速フーリエ変換されて周波数軸上の信号に変換され、周波数シフト回路235で高域側に周波数シフト処理され、逆FFT回路236で逆高速フーリエ変換されることにより、高域側の時間軸信号とされ、重畳加算回路237を介して加算器228に送られる。

【0051】加算器228では、上記アップサンプリング回路227からの信号と加算され、出力端子229より16 kbpsのビットレートの一部に対応する音声信号として取り出される。全体としての16 kbpsのビットレートの信号は、上記出力端子219からの信号も合成されることにより取り出される。

【0052】ここで、スケラビリティについて説明する。上記図1、図2の構成においては、6 kbpsと16 kbpsとの2通りの伝送ビットレートをほぼ同様な符号化復号化方式で実現しており、16 kbpsのビットストリーム内に6 kbpsのビットストリームを完全に包含するスケラビリティを実現しているが、さらに2 kbpsのような極端にビットレートの異なる符号化復号化を行う場合には、このような完全な包含関係を得るのは難しい。

【0053】ここで、同一の符号化復号化方式を適用できない場合であっても、最大限に共有関係を持ちながらスケラビリティを持たせることが好ましい。

【0054】このため、図3に示すような構成の符号化

装置によって2 kbps の符号化を行い、図1の構成との間に最大限の共有部分あるいは共有データを持たせ、全体として16 kbps のビットストリームで、この内16 kbps 全てを使用する場合と、6 kbps を使用する場合と、2 kbps を移用する場合とを、それぞれ用途に応じて使い分けようとしている。

【0055】なお、厳密には、後述するように、2 kbps では2 kbps の情報を全て使用するが、6 kbps のモードでは、符号化単位となるフレームが有声音(V)のとき6 kbps、無声音(UV)のとき5.65 kbps であり、また16 kbps のモードでは、フレームが有声音(V)のとき15.2 kbps、無声音(UV)のとき14.85 kbps である。

【0056】ここで、図3の2 kbps の符号化装置の構成及び動作を説明する。

【0057】図3に示す符号化装置の基本的な考え方は、入力音声信号の短期予測残差例えばLPC(線形予測符号化)残差を求めてサイン波分析(sinusoidal analysis)符号化、例えばハーモニックコーディング(harmonic coding)を行う第1の符号化部310と、入力音声信号に対して位相伝送を行う波形符号化により符号化する第2の符号化部320とを有し、入力信号の有声音(V:Voiced)の部分の符号化に第1の符号化部310を用い、入力信号の無声音(UV:Unvoiced)の部分の符号化には第2の符号化部320を用いるようにすることである。

【0058】上記第1の符号化部310には、例えばLPC残差をハーモニック符号化やマルチバンド励起(MBE)符号化のようなサイン波分析符号化を行う構成が用いられる。上記第2の符号化部320には、例えば合成による分析法を用いて最適ベクトルのクローズドループサーチによるベクトル量子化を用いた符号励起線形予測(CELP)符号化の構成が用いられる。

【0059】図3の例では、入力端子301に供給された音声信号が、第1の符号化部310のLPC逆フィルタ311及びLPC分析・量子化部313に送られている。LPC分析・量子化部313から得られたLPC係数あるいはいわゆる $\alpha$ パラメータは、LPC逆フィルタ311に送られて、このLPC逆フィルタ311により入力音声信号の線形予測残差(LPC残差)が取り出される。また、LPC分析・量子化部313からは、後述するようにLSP(線スペクトル対)の量子化出力が取り出され、これが出力端子302に送られる。LPC逆フィルタ311からのLPC残差は、サイン波分析符号化部314に送られる。サイン波分析符号化部314では、ピッチ検出やスペクトルエンベロープ振幅計算が行われると共に、V(有声音)/UV(無声音)判定部315によりV/UVの判定が行われる。サイン波分析符号化部314からのスペクトルエンベロープ振幅データがベクトル量子化部316に送られる。スペクトルエン

ベロープのベクトル量子化出力としてのベクトル量子化部316からのコードブックインデックスは、スイッチ317を介して出力端子303に送られ、サイン波分析符号化部314からの出力は、スイッチ318を介して出力端子304に送られる。また、V/UV判定部315からのV/UV判定出力は、出力端子305に送られると共に、スイッチ317、318の制御信号として送られており、上述した有声音(V)のとき上記インデックス及びピッチが選択されて各出力端子303及び304からそれぞれ取り出される。

【0060】図3の第2の符号化部320は、この例ではCELP(符号励起線形予測)符号化構成を有しており、雑音符号帳321からの出力を、重み付きの合成フィルタ322により合成処理し、得られた重み付き音声信号を減算器323に送り、入力端子301に供給された音声信号を聴覚重み付けフィルタ325を介して得られた音声との誤差を取り出し、この誤差を距離計算回路324に送って距離計算を行い、誤差が最小となるようなベクトルを雑音符号帳321でサーチするような、合成による分析(Analysis by Synthesis)法を用いたクローズドループサーチを用いた時間軸波形のベクトル量子化を行っている。このCELP符号化は、上述したように無声音部分の符号化に用いられており、雑音符号帳321からのUVデータとしてのコードブックインデックスは、上記V/UV判定部315からのV/UV判定結果が無声音(UV)のときオンとなるスイッチ327を介して、出力端子307より取り出される。

【0061】このような符号化装置のLPC分析・量子化部313が図1のLPC分析・量子化部130の一部として共用でき、端子302からの出力がそのまま図1の出力端子131からの出力として使用できる。また、サイン波分析符号化部314により得られるピッチデータの一部が図1のピッチ分析回路115からの出力として使用でき、このピッチ分析回路115をサイン波分析符号化部314内のピッチ出力部分と共用することも可能である。

【0062】このように、図3の符号化方式と図1の符号化方式とは異なっているが、両者とも共通する情報を持っており、図4に示すようなスケラビリティを有している。

【0063】この図4において、2 kbps のビットストリームS2は、分析合成フレームがV(有声音)のときとUV(無声音)のときとで内部構造が異なっており、Vのときの2 kbps のビットストリームS2vは、2つの部分S2ve、S2vaから、UVのときの2 kbps のビットストリームS2uは、2つの部分S2ue、S2uaからそれぞれ成っている。部分S2veは、ピッチラグが1フレーム160サンプル当たり1ビット(以下、1ビット/160サンプルのように示す)で、振幅Amが15ビット/160サンプルであり、計16ビット/160サンプルとなる。これは、8 kHz

サンプリングで0.8 kbps のビットレートのデータに相当する。部分S 2ueは、LPC残差が $11\text{ビット}/80\text{サンプル}$ と、予備の $1\text{ビット}/160\text{サンプル}$ とで、計 $23\text{ビット}/160\text{サンプル}$ となり、1.15 kbps のビットレートのデータに相当する。2 kbps のビットストリームS 2の残りの部分S 2va、S 2uaは、上述した6 kbps、16 kbps との共有部分あるいは共通部分であり、部分S 2vaは、LSPデータ $32\text{ビット}/320\text{サンプル}$ と、V/U判定データ $1\text{ビット}/160\text{サンプル}$ と、ピッチラグ $7\text{ビット}/160\text{サンプル}$ とで、計 $24\text{ビット}/160\text{サンプル}$ となり、1.2 kbps のビットレートのデータに相当する。部分S 2uaは、LSPデータ $32\text{ビット}/320\text{サンプル}$ と、V/U判定データ $1\text{ビット}/160\text{サンプル}$ とで、計 $17\text{ビット}/160\text{サンプル}$ となり、0.85 kbps のビットレートのデータに相当する。

【0064】また、6 kbps のビットストリームS 6は、上記ビットストリームS 2と同様に、分析フレームがVのときとUVのときとで内部構造が一部だけ異なる。Vのときの6 kbps のビットストリームS 6vは、2つの部分S 6va、S 6vbから、UVのときの6 kbps のビットストリームS 6uは、2つの部分S 6ua、S 6ubからそれぞれ成っている。部分S 6vaは、上述したように、部分S 2vaと共通のデータ内容であり、部分S 6vbは、ピッチゲイン $6\text{ビット}/160\text{サンプル}$ と、ピッチ残差 $18\text{ビット}/32\text{サンプル}$ とで、計 $96\text{ビット}/160\text{サンプル}$ となり、4.8 kbps のビットレートのデータに相当する。また、部分S 6uaは、上記部分S 2uaと共通のデータ内容であり、部分S 6ubは、上記部分S 6vbと共通のデータ内容である。

【0065】また、16 kbps のビットストリームS 16は、上記ビットストリームS 2及びS 6と同様に分析フレームがVのときとUVのときとで内部構造が一部だけ異なる。Vのときの16 kbps のビットストリームS 16vは、4つの部分S 16va、S 16vb、S 16vc、S 16vdから、UVのときの16 kbps のビットストリームS 16uは、4つの部分S 16ua、S 16ub、S 16uc、S 16udからそれぞれ成っている。部分S 16vaは、上記部分S 2va、S 6vaと共通のデータ内容であり、S 16vbは、上記部分S 6vb、S 6ubと共通のデータ内容である。部分S 16vcは、ピッチラグ $2\text{ビット}/160\text{サンプル}$ と、ピッチゲイン $11\text{ビット}/160\text{サンプル}$ と、ピッチ残差 $18\text{ビット}/32\text{サンプル}$ と、S/Mモードデータ $1\text{ビット}/160\text{サンプル}$ とで、計 $104\text{ビット}/160\text{サンプル}$ となり、5.2 kbps のビットレートに相当する。なお、上記S/Mモードデータは、VQ回路124で、音声(Speech)用と楽音(Music)用とで異なる2種類のコードブック(符号帳)を切り換えるためのものである。部分S 16vdは、高域LPCデータ $5\text{ビット}/160\text{サンプル}$ と、高域LPC残差 $15\text{ビット}/32\text{サンプル}$ とで、計 $80\text{ビット}/160\text{サンプル}$ となり、4 kbps のビットレートに相当する。また、部分S 16uaは、上記部分S 2ua、S 6uaと共通のデータ内容であり、部分S 16

ubは、上記部分S 16vbすなわち上記部分S 6vb、S 6ubと共通のデータ内容である。さらに、部分S 16ucは、上記部分S 16vcと共通のデータ内容であり、部分S 16udは、上記部分S 16vdと共通のデータ内容である。

【0066】以上のようなビットストリームを得るための図1、図3の構成をまとめると、図5のようになる。

【0067】この図5において、入力端子11は図1、図3の入力端子101に対応し、これが図1のLPF102、サンプリング周波数変換器103、減算器106、BPF107等に相当する帯域分割回路12に送られて、低域側と高域側とに分割される。帯域分割回路12からの低域側信号は、図3の構成に対応する2 k符号化部21と共通部分符号化部22とに送られる。共通部分符号化部22は、図1のLPC分析・量子化部130、あるいは図3のLPC分析・量子化部310にほぼ相当し、さらに図3のサイン波分析符号化部内のピッチ抽出部分や図1のピッチ分析回路115も共通部分符号化部22に含ませることもできる。

【0068】また、帯域分割回路12からの低域側信号は、6 k符号化部23及び12 k符号化部24にも送られる。6 k符号化部23は、図1の回路111～116にほぼ相当し、12 k符号化部は、図1の回路117、118、122～128にほぼ相当する。

【0069】帯域分割回路12からの高域側信号は、高域4 k符号化部25に送られる。高域4 k符号化部25は、図1の回路161～164、171、172にほぼ相当する。

【0070】この図5の各出力端子31～35から出力されるビットストリームと図4の各部分との関係を説明する。2 k符号化部21から出力端子31を介して、図4の部分S 2ve又はS 2ueのデータが出力され、共通部分符号化部22から出力端子32を介して、図4の部分S 2va(=S 6va=S 16va)又はS 2ua(=S 6ua=S 16ua)のデータが出力される。また、6 k符号化部23から出力端子33を介して、図4の部分S 6vb(=S 16vb)又はS 6ub(=S 16ub)のデータが出力される。さらに、12 k符号化部24から出力端子34を介して、図4の部分S 16vc又はS 16ucのデータが出力され、高域4 k符号化部25から出力端子35を介して、図4の部分S 16vd又はS 16udのデータが出力される。

【0071】以上説明したスケーラビリティの実現の技術を一般化すると、入力信号に対して第1の符号化を施して得られた第1の符号化信号と、上記入力信号に対して上記第1の符号化の一部とのみ共通する部分と共通しない部分とを有し上記第1の符号化とは独立の第2の符号化を施して得られた第2の符号化信号とを多重化する際に、上記第1の符号化信号と、上記第2の符号化信号の内の上記第1の符号化信号と共通する部分を除く信号

とを多重化することである。

【0072】これによって、本質的に異なる符号化方式であっても、共有できるものを最大限に共有させて、スケラビリティを持たせることができる。

【0073】次に、上記図1、図2の各部のより具体的な動作について説明する。

【0074】先ず、図6の(A)に示すように、フレーム間隔、いわゆるフレームインターバルをNサンプル、例えば160サンプルとし、1フレームに1回の分析を行う場合について説明する。

【0075】ピッチ分析中心を  $t = kN$  (ただし  $k=0, 1, 2, 3, \dots$ ) とするとき、LPC逆フィルタ111からのLPC予測残差について、 $t = kN - N/2 \sim kN + N/2$  に存在する成分から成る次元数Nのベクトルを  $\underline{X}$  とし、これをLサンプルだけ時間軸の前方にずらした  $t = kN - N/2 - L \sim kN + N/2 - L$  の成分から成るN次元ベクトルを  $\underline{X}_L$  として、

$$\|\underline{X} - g \underline{X}_L\|^2$$

が最小となるよう  $L = L_{opt}$  をサーチし、この  $L_{opt}$  をこの区間での最適ピッチラグ  $L_1$  とする。あるいは、ピッチの急激な変化を避けるため、ピッチトラッキングを行った後の値を最適ピッチラグ  $L_1$  としてもよい。

【0076】次に、この最適ピッチラグ  $L_1$  に対して、

【0077】

【数1】

$$\begin{aligned} L((k-1/2)N) &= L(kN) \\ &= (L(kN) + L((k-1)N)) / 2 \\ &= L((k-1)N) \end{aligned}$$

これらの内のどの値を採用するかは、それぞれのラグに対応したピッチ残差のパワーを計算することによってなされる。

【0084】すなわち、 $t = (k-1/2)N$  を中心とした  $t = (k-1/2)N - N/4 \sim (k-1/2)N + N/4$  の次元数  $N/2$  のベクトルを  $\underline{X}$  とし、 $L(kN)$ 、 $(L(kN) + L((k-1)N))/2$ 、 $L((k-1)N)$  だけそれぞれ遅れた次元数  $N/2$  のベクトルを  $\underline{X}_0^{(0)}$ 、 $\underline{X}_1^{(0)}$ 、 $\underline{X}_2^{(0)}$  とし、これらの各ベクトル  $\underline{X}_0^{(0)}$ 、 $\underline{X}_1^{(0)}$ 、 $\underline{X}_2^{(0)}$  及び  $\underline{X}_2^{(0)}$  のそれぞれの近傍のベクトルを  $\underline{X}_0^{(-1)}$ 、 $\underline{X}_0^{(1)}$ 、 $\underline{X}_1^{(-1)}$ 、 $\underline{X}_1^{(1)}$  及び  $\underline{X}_2^{(-1)}$ 、 $\underline{X}_2^{(1)}$  とする。また、これらの各ベクトル  $\underline{X}_0^{(i)}$ 、 $\underline{X}_1^{(i)}$ 、 $\underline{X}_2^{(i)}$  (ただし  $i = -1, 0, 1$ ) に対応する核ピッチゲイン  $g_0^{(i)}$ 、 $g_1^{(i)}$ 、 $g_2^{(i)}$  について、

【0085】

【数3】

$$D = \left\| \underline{X} - \sum_{i=-1}^1 g_i \underline{X}_{L_1+i} \right\|^2$$

【0078】が最小となる  $g_i$  の組を

【0079】

【数2】

$$\frac{\partial D}{\partial g_i} = 0 \quad (\text{ただし, } i = -1, 0, 1)$$

10 【0080】について解き、ピッチゲインベクトル  $\underline{g}_1$  を求める。このピッチゲインベクトル  $\underline{g}_1$  をベクトル量子化したもののコードブックインデックスを  $g_1$  とする。

【0081】次に、さらに予測精度を上げるため、 $t = (k-1/2)N$  にも分析中心をおくことを考える。このとき、予め  $t = kN$  及び  $(k-1)N$  でのピッチラグ、ピッチゲインがそれぞれ求められているものとする。

【0082】音声信号の場合、その基本周波数はゆるやかに変化すると考えられるため、 $t = kN$  のときのピッチラグ  $L(kN)$  と、 $t = (k-1)N$  のときのピッチラグ  $L((k-1)N)$  ととの間に大きな変化はないと考えられ、またその変化も線形であると考えられるため、 $t = (k-1/2)N$  のときのピッチラグ  $L((k-1/2)N)$  のとり得る値に制限を加えることは可能である。本例では、次のようにしている。

【0083】

$$D_0 = \left\| \underline{X} - \sum_i g_0^{(i)} \underline{X}_0^{(i)} \right\|^2$$

$$D_1 = \left\| \underline{X} - \sum_i g_1^{(i)} \underline{X}_1^{(i)} \right\|^2$$

$$D_2 = \left\| \underline{X} - \sum_i g_2^{(i)} \underline{X}_2^{(i)} \right\|^2$$

【0086】の3つうちの最小のもの  $D_j$  に対するラグを、 $t = (k-1/2)N$  での最適ラグ  $L_2$  とし、そのときのピッチゲイン  $g_j^{(i)}$  (ただし  $i = -1, 0, 1$ ) をベクトル量子化した上でピッチゲインを求める。なお、 $L_2$  のとり得る値は3通りであり、これは現在及び過去の  $L_1$  から求められるため、ストレートな値ではなく補間スキームを表すラグを補間インデックスとして伝送すればよい。また、 $L(kN)$ 、 $L((k-1)N)$  のいずれかが0、すなわちピッチが無い、ピッチ予測利得がとれない、と判断されるときには、 $L((k-1/2)N)$  の候補として上記  $(L(kN) + L((k-1)N))/2$  は除外される。

【0087】このように、ピッチラグ算出に用いるベクトル  $\underline{X}$  の次元数を半分の  $N/2$  にした場合、 $t = kN$  が分析中心のときの  $L_k$  はそのまま用いることができるが、 $\underline{X}$

の次元数がNで分析したときのピッチゲインが得られて  
いるにも拘わらず、再度ゲイン計算を行い、そのデータ  
を伝送しなければならない。ここでは、そのビット数削

$$\hat{g}_{ld} = \hat{g}_l - \hat{g}_i \text{ を量子化する}$$

ただし、分析長 = N で求めた時の量子化済のピッチゲイン (ベクトル)  $\hat{g}_l$

$$\text{分析長} = \frac{N}{2} \text{ で求めた時の量子化していないピッチゲイン } \hat{g}_i$$

【0089】ベクトル  $\underline{g}$  の要素 ( $g_0, g_1, g_2$ ) の内、  
 $g_1$  がもっとも大きく  $g_0, g_2$  は0に近い、あるいは  
その逆であり、ベクトル  $\underline{g}$  は3点の間で強い相関がある  
ので、上記ベクトル  $\underline{g}_{ld}$  は元のベクトル  $\underline{g}$  に比べて分散  
が小さくなることが予想され、より少ないビット数で量  
子化できる。

【0090】従って、1フレームで伝送すべきピッチパ  
ラメータは、 $L_1, g_1, L_2, g_2, g_{ld}$  の5つになる。

【0091】次に、図5の(B)は、フレーム周波数の  
8倍のレートで補間されたLPC係数の位相を示してお  
り、このLPC係数は、図1のLPC逆フィルタ111  
による予測残差算出に用いられ、また図2のLPC合成  
フィルタ215、225、ピッチスペクトラルポストフ  
ィルタ216、226にそれぞれ用いられる。

【0092】次に、上記ピッチラグ及びピッチゲインか  
ら求められたピッチ残差のベクトル量子化について説明  
する。

【0093】ベクトル量子化の聴覚重み付けを容易にま  
た精度よく行うため、ピッチ残差は50%オーバーラ  
ップの窓かけをした上、MDCT変換を行い、この領域で  
重み付けベクトル量子化を行う。このときの変換長は任  
意であるが、以下の点を考慮して上で、この例ではかな  
り小次元なものをを用いている。

【0094】(1) 大次元のベクトル量子化は演算量が膨  
大になり、MDCT領域でスプリットあるいは並べ替え  
を行わざるを得ない。

(2) スプリットした場合、スプリットされたバンド間の  
ビットアロケーションを精密に行うことは大変困難であ  
る。

(3) 次元数が2のべき乗ではないとき、FFTを用いた  
MDCTの高速算法が使用できない。

【0095】今回はフレーム長を20 msec (= 160  
サンプル/8 kHz) にとっているため、160/5 = 3  
2 = 2<sup>5</sup> となることから、50%オーバーラップを考慮  
してMDCT変換サイズを64にとり、上記(1)~(3)の  
各点の解決を図った。

【0096】フレーミングの状態は図6の(C)のよう  
になる。

【0097】すなわち、この図6の(C)において、2  
0 msec = 160 サンプルのフレーム内のピッチ残差  $r_p(n)$   
(ただし  $n=0, 1, \dots, 191$ 、ここで  $n=160, \dots, 191$  は、

減のため、

【0088】

【数4】

10 次のフレームの0, ..., 31の意味) を5つのサブフレーム  
に分け、5つのサブフレームのi番目 ( $i=0, 1, \dots, 4$ ) の  
サブフレームのピッチ残差  $r_{pi}(n)$  (ただし、 $n=0, 1,$   
..., 31) を、

$$r_{pi}(n) = r_p(32i+n)$$

とする。このサブフレームのピッチ残差  $r_{pi}(n)$  に、M  
DCTのエリアシング相殺ができるような窓関数  $w(n)$   
をかけて得られる  $w(n) \cdot r_{pi}(n)$  に対してMDCT変換  
を施す。この窓関数  $w(n)$  としては、例えば

$$w(n) = \sqrt{(1 - (\cos 2\pi(n+0.5))/64))}$$

を用いればよい。

【0098】なお、MDCTの変換処理の演算は、変換  
長が64 (= 2<sup>6</sup>) のため、FFTを用いて次のように  
計算できる。

【0099】(1)  $x(n) = w(n) \cdot r_{pi}(n) \cdot \exp((-2\pi j/64)(n/2))$  とする。

(2)  $x(n)$  を64ポイントFFT処理し、これを  $y(k)$   
とする。

(3)  $y(k) \cdot \exp((-2\pi j/64)(k+1/2)(1/2+64/4))$  の実部  
を取り、これをMDCT係数  $c_i(k)$  (ただし、 $k=0, 1,$   
..., 31) とする。

【0100】次に、各サブフレームのMDCT係数  $c_i(k)$   
をベクトル量子化するが、このときの重み付けにつ  
いて説明する。

【0101】ピッチ残差  $r_{pi}(n)$  をベクトル  $\underline{r}_i$  とお  
くと、合成後の距離Dは、

【0102】

【数5】

$$\begin{aligned} D^2 &= \|H(\underline{r}_i - \hat{\underline{r}}_i)\|^2 \\ &= (\underline{r}_i - \hat{\underline{r}}_i)' H' H (\underline{r}_i - \hat{\underline{r}}_i) \\ &= (\underline{r}_i - \hat{\underline{r}}_i)' M' M H' H M (\underline{r}_i - \hat{\underline{r}}_i) \\ &= (\underline{c}_i - \hat{\underline{c}}_i)' M H' H M (\underline{c}_i - \hat{\underline{c}}_i) \end{aligned}$$

(ただし、 $H$  は合成フィルタ行列、 $M$  はMDCT行列、  
 $\underline{c}_i$  は  $c_i(k)$  のベクトル表記、  
 $\hat{\underline{c}}_i$  は  $\hat{c}_i(k)$  を量子化した  $\hat{c}_i(k)$  のベクトル表現)

【0103】ここで、 $M$  は、その性質から  $H^t H$  (た  
だし  $H^t$  は  $H$  の転置行列) を対角化すると考えられるの

50 で、

【0104】

【数6】

$$MH^T HM = \begin{bmatrix} h_0^2 & & & \\ & h_1^2 & 0 & \\ & 0 & \ddots & \\ & & & h_{n-1}^2 \end{bmatrix} \quad (n=64)$$

【0105】とし、ここでは $h_i$ を合成フィルタの周波数応答にとった。従って、

【0106】

【数7】

$$D^2 = \sum_k h_k^2 (c_i(k) - \hat{c}_i(k))^2$$

【0107】このように、 $h_k$ をそのまま $c_i(k)$ の量子化の重み付けに使用した場合、合成後のノイズがフラットになる、いわゆる100%ノイズシェイピングになるため、さらに聴覚重み付け $W$ によりフォルマントを相似形のノイズになるようにコントロールする。

【0108】

【数8】

$$W = \begin{bmatrix} w_0 & & & \\ & w_1 & 0 & \\ & 0 & \ddots & \\ & & & w_{n-1} \end{bmatrix} \quad (n=64)$$

【0109】なお、 $h_i^2$ 、 $w_i^2$ は、合成フィルタ $H(z)$ 及び聴覚重み付けフィルタ $W(z)$

【0110】

【数9】

$$H(z) = \frac{1}{1 + \sum_{j=1}^P \alpha_{ij} z^{-j}}$$

$$W(z) = \frac{1 + \sum_{j=1}^P \lambda_b^j \alpha_{ij} z^{-j}}{1 + \sum_{j=1}^P \lambda_a^j \alpha_{ij} z^{-j}}$$

( $P$ は分析次数、 $\lambda_a$ 、 $\lambda_b$ は重み付けのための係数)

【0111】のインパルス応答のFFTパワースペクトルとして求められる。

【0112】ここで、 $\alpha_{ij}$ は、第 $i$ サブフレームに対応するLPC係数であり、補間されたLSP係数から求められる。すなわち、前フレームの分析で得られた $LSP_0(j)$ と現フレームの $LSP_1(j)$ とを内分し、本例の場合、第 $i$ サブフレームのLSPは、

【0113】

【数10】

$$LSP^{(i)}(j) = \left(1 - \frac{i+1}{5}\right) LSP_0(j) + \frac{i+1}{5} LSP_1(j)$$

(ただし、 $i=0,1,2,3,4$ )

【0114】として $LSP^{(i)}(j)$ を求める。その後、 $LSP \rightarrow \alpha$ 変換により $\alpha_{ij}$ を求める。

【0115】このようにして求められた $H$ 、 $W$ に対して、新たに $W' = WH$ と置き、ベクトル量子化の際の距離尺度として用いる。

【0116】ベクトル量子化は、シェイプ、ゲインベクトル量子化によって行うが、その学習時の最適エンコード、デコード条件について説明する。

【0117】学習のある時点でのシェイプコードブックを $\underline{s}$ 、ゲインコードブックを $\underline{g}$ とし、トレーニング時の入力すなわち各サブフレームでのMDCT係数を $\underline{x}$ 、そのサブフレームでの重みを $W'$ とすると、このときの歪のパワー $D^2$ は、以下の式で定義される。

$$D^2 = \|W'(\underline{x} - \underline{g}\underline{s})\|^2$$

この $D^2$ を最小にするような $(\underline{g}, \underline{s})$ を選択することが最適エンコード条件である。

【0119】

【数11】

$$D^2 = (\underline{x} - \underline{g}\underline{s})^T W'^T W' (\underline{x} - \underline{g}\underline{s})$$

$$= \underline{s}^T W'^T W' \underline{s} \left( \underline{g} - \frac{\underline{s}^T W'^T W' \underline{x}}{\underline{s}^T W'^T W' \underline{s}} \right)^2 + \underline{x}^T W'^T W' \underline{x} - \frac{(\underline{s}^T W'^T W' \underline{x})^2}{\underline{s}^T W'^T W' \underline{s}}$$

【0120】したがって、まず第1のステップとして、シェイプコードブックについて、

【0121】

【数12】

$$\frac{(\underline{s}^T W'^T W' \underline{x})^2}{\underline{s}^T W'^T W' \underline{s}}$$

【0122】を最大にする $\underline{s}_{opt}$ をサーチし、ゲインコードブックについては、この $\underline{s}_{opt}$ に対し、

【0123】

【数13】

$$\frac{\underline{s}_{opt}^T W'^T W' \underline{x}}{\underline{s}_{opt}^T W'^T W' \underline{s}_{opt}}$$

【0124】に最も近い $\underline{g}_{opt}$ をサーチすればよい。

【0125】次に、最適デコード条件を求める。

【0126】第2のステップとしてシェイプコードブックについて、学習中のある時点でシェイプコードブック $\underline{s}$ にエンコードされた $\underline{x}$ の集合 $\underline{x}_k$  ( $k=0, \dots, N-1$ )に対して、このときの歪の総和 $E_s$ は、

【0127】

【数14】

$$E_s = \sum_{k=0}^{N-1} \|w'_k(x_k - g_k s)\|^2$$

【0128】であるから、これを最小にする  $s$  は、

【0129】

【数15】

$$\frac{\partial E_s}{\partial s} = 0 \text{ より}$$

【0130】より

【0131】

【数16】

$$s = \left( \sum_{k=0}^{N-1} g_k^2 w'_k w'_k \right)^{-1} \sum_{k=0}^{N-1} g_k w'_k w'_k x_k$$

【0132】と求められる。

【0133】ゲインコードブックについては、ゲインコードブック  $g$  にエンコードされた  $\underline{x}$  の集合  $\underline{x}_k$  (重み  $W'_k$ 、シェイプ  $s_k$ ) について、歪の総和  $E_g$  は、

【0134】

【数17】

$$E_g = \sum_{k=0}^{N-1} \|w'_k(x_k - g s_k)\|^2$$

$$\frac{\partial E}{\partial g} = 0 \text{ より}$$

$$g = \frac{\sum_{k=0}^{M-1} s'_k w'_k w'_k x_k}{\sum_{k=0}^{M-1} s'_k w'_k w'_k s_k}$$

【0135】である。

【0136】上記第1、第2のステップを繰り返し求めながら、GLA (一般化ロイドアルゴリズム) によって、シェイプ、ゲインコードブックを得ることができる。

【0137】なお、本例では、信号レベルの小さいときのノイズを重視するため、 $W'$  そのものでなく、レベル (の逆数) の重みをつけた  $W' / \|\underline{x}\|$  を用いて学習を行っている。

【0138】このようにして、作成した符号帳を用いてピッチ残差をMDCTしたものに対するベクトル量子化を行い、そのインデクスをLPC (実際にはLSP)、ピッチ、ピッチゲインとともに伝送し、デコード側では逆ベクトル量子化、ピッチ、LPC合成を行うことで、再生音を得ることができるが、本例では、さらにレートの高い動作を可能とするため、前述のピッチラグ、ピッチゲイン算出の頻度向上とともに、ピッチ残差MD

CTベクトル量子化を多段にすることで、これに対応している。

【0139】一例を図7の(A)に示す。ここでの段数は2段であり、シーケンシャルな多段ベクトル量子化であるが、2段目の入力1段目のデコード結果を  $L_2$ 、 $g_2$ 、 $g_{ld}$  から作られた精度の高いピッチ残差から引いたものとして用いる。すなわち、1段目のMDCT回路113からの出力をVQ回路114でベクトル量子化した後の代表ベクトルあるいは逆量子化出力を、逆MDCT回路113aで逆MDCT処理した結果を、減算器128'に送り、2段目の残差 (図1のピッチ逆フィルタ122からの出力) から減算している。この減算器128'からの出力をMDCT回路123'によりMDCT処理してVQ回路124で量子化する。これは、1段目の逆MDCTを行わない等価な図7の(B)のような構成とすることができ、図1ではこの(B)の構成を用いている。

【0140】図2のデコーダ側でMDCT係数のインデクス  $I_{dx}V_{q1}$ 、 $I_{dx}V_{q2}$  をともに用いたデコードをする際は、 $I_{dx}V_{q1}$ 、 $I_{dx}V_{q2}$  の逆ベクトル量子化の結果の和を逆MDCT、オーバーラップ加算をした上で、ピッチ合成、LPC合成を行い再生音を得る。当然ピッチ合成時のピッチラグ、ピッチゲイン更新頻度は、1段のみの場合の倍になり、本願では、80サンプル毎に切り換わるピッチ合成フィルタを駆動することになる。

【0141】次に、図2のデコーダ側のポストフィルタ216、226について説明する。

【0142】ポストフィルタ216、226は、ピッチ強調、高域強調、スペクトル強調フィルタの縦続接続でポストフィルタ特性  $p(z)$  を実現する。

【0143】

【数18】

$$P(z) = \frac{1}{1 - \gamma_p \sum_{i=1}^p g_i z^{-L+1}} (1 - \gamma_b z^{-1}) \frac{1 - \sum_{j=1}^p \gamma_n^j \alpha_{ij} z^{-j}}{1 - \sum_{j=1}^p \gamma_d^j \alpha_{ij} z^{-j}}$$

【0144】この式において、 $g_i$ 、 $L$  はピッチ予測で求められたピッチゲイン、ピッチラグであり、 $\gamma$  はピッチ強調の度合いを表すパラメータである (例えば  $\gamma = 0.5$ )。また、 $\gamma_b$  は高域強調 (例えば  $\gamma_b = 0.4$ ) を、 $\gamma_n$ 、 $\gamma_d$  はスペクトルの強調度 (例えば  $\gamma_n = 0.5$ 、 $\gamma_d = 0.8$ ) を表すパラメータである。

【0145】次に、LPC合成フィルタの出力  $s(n)$ 、ポストフィルタの出力  $s_p(n)$  についてのゲイン補正を行う。このときの係数  $k_{adj}$  は、

【0146】

【数19】

$$k_{adj} = \frac{\sum_{n=0}^{N-1} (s(n))^2}{\sum_{n=0}^{N-1} (s_p(n))^2} \quad (N=80 \text{ または } 160)$$

【0147】であるが、 $k_{adj}$  はフレーム内で固定ではなく、LPFを通した上でサンプル毎に変化させる。なお、 $p$ として例えば0.1が用いられる。

【0148】 $k_{adj}(n) = (1-p)k_{adj}(n-1) + pk_{adj}$   
次に、フレームのつながりを滑らかにするため、以下のよう  
10 にピッチ強調フィルタを2つ用意し、その結果をクロス  
フェイドしたものを最終出力とする。

【0149】

【数20】

$$\frac{1}{1 - \gamma_p \sum_{i=1}^L g_{0i} z^{-L+i}}$$

【0150】

【数21】

$$\frac{1}{1 - \gamma_p \sum_{i=1}^L g_i z^{-L+i}}$$

【0151】これにより構成されたポストフィルタの出力  $s_{p0}(n)$ 、 $s_p(n)$  に対して最終出力  $s_{out}(n)$  を、  
 $s_{out}(n) = (1-f(n)) s_{p0}(n) + f(n) s_p(n)$   
とする。ここで、 $f(n)$  は、例えば図8に示すような窓  
である。この図8の(A)は低レート時、(B)は高レ  
ート時をそれぞれ示しており、(B)の80サンプル幅  
30 の窓は、160サンプル、20 msec の合成時には2回  
繰り返して用いられる。

【0152】次に、図1のエンコーダ側のVQ(ベクトル量子化)回路124について説明する。

【0153】このVQ回路124は、音声(Speech)用と楽音(Music)用とで互いに異なる2種類の符号帳(コードブック)を有しており、これらの2種類の符号帳を入力信号に応じて切り換え選択するようにしている。

【0154】すなわち、音声、楽音信号の量子化において、量子化器の構成が決まっている場合、この量子化器の持つ符号帳は、学習時に使用した音声、楽音の性質において最適なものとなるため、両者を一緒にして学習した場合、両者の性質が大きく異なると、学習後の符号帳は両者の平均的な性質を持つことになる。従って、一つの符号帳で量子化器を構成した場合、そのパフォーマンス、あるいは平均S/Nは、あまり高くないことが予想される。

【0155】そこで、本例においては、このように性質の異なる複数の信号について、それぞれの学習データを

用いて作成した符号量を切り換えて、量子化器の性能を向上させている。

【0156】図9は、このような2種類の符号帳  $C_{BA}$ 、 $C_{BB}$  を有するベクトル量子化器の概略的な構成を示している。

【0157】この図9において、入力端子501に供給された入力信号は、ベクトル量子化器511、512に送られる。これらのベクトル量子化器511、512は、それぞれ符号長(コードブック)  $C_{BA}$ 、 $C_{BB}$  を有している。これらのベクトル量子化器511、512からの代表ベクトルあるいは逆量子化出力は、それぞれ減算器513、514に送られ、元の入力信号との差がとられて、これらの各誤差分が比較器515に送られる。比較器515では、各誤差分を比較して、誤差が小さい方のベクトル量子化器511、512からの量子化出力であるインデックスを切換スイッチ516で切換選択して出力端子502に送る。

【0158】この場合、各ベクトル量子化器511、512の量子化単位時間あるいは周期よりも、切換スイッチ516の切換周期を長くしている。例えば、量子化単位がフレームを8分割したサブフレームであるとき、切  
20 換スイッチ516をフレーム単位で切り換えている。

【0159】ここで、例として、それぞれ音声のみ、楽音のみで学習した同じサイズNで、同じ次元Mの符号帳  $C_{BA}$ 、 $C_{BB}$  があるとして、あるフレームのL個のデータから成るL次元のデータ  $\underline{X}$  をサブフレーム長M(=L/n)でベクトル量子化したとき、量子化後の歪みについて、符号帳  $C_{BA}$  を用いたときを  $E_A(k)$ 、符号帳  $C_{BB}$  を用いたときを  $E_B(k)$  とする。これらの歪み  $E_A(k)$ 、  
30  $E_B(k)$  は、それぞれインデックス  $i$ 、 $j$  が選ばれたとして、

$$E_A(k) = \| W_k (\underline{X} - \underline{C}_{Ai}) \|^2$$

$$E_B(k) = \| W_k (\underline{X} - \underline{C}_{Bj}) \|^2$$

である。この式で、 $W_k$  はサブフレーム  $k$  での重み付け行列を表し、 $\underline{C}_{Ai}$ 、 $\underline{C}_{Bj}$  は符号帳  $C_{BA}$ 、 $C_{BB}$  のそれぞれインデックス  $i$ 、 $j$  に対応する代表ベクトルを表す。

【0160】このようにして得られた2つの歪みに対して、1フレーム内での歪みの総和により、そのフレームに最適な符号帳を採用することを考える。このときの選  
40 び方について、次の2つの方法が考えられる。

【0161】第1の方法として、全てのサブフレームにおいて、符号帳  $C_{BA}$  及び  $C_{BB}$  のみを用いて量子化を行い、歪みのフレーム内総和  $\sum_k E_A(k)$ 、 $\sum_k E_B(k)$  を求め、小さい方の歪みの総和を与える符号帳  $C_{BA}$ 、 $C_{BB}$  のいずれかを1フレームに亘って使用する。

【0162】この第1の方法を実現する構成例を図10に示す。この図10では、上記図9と対応する部分に同じ参照番号を付しており、参照番号に添付したa、b、  
50 …等の添字は、サブフレーム  $k$  に対応している。符号帳  $C_{BA}$  については、サブフレーム毎の歪みが得られ



る各減算器 513a、513b、・・・、513nからの出力のフレーム内総和を加算器 517でとり、符号帳 CBB については各サブフレーム毎の歪みのフレーム内総和を加算器 518でとって、これらを比較器 515で比較することにより、コードブック切換のための制御信号あるいは選択フラグを端子 503より得ている。

【0163】次に、第2の方法は、各サブフレーム毎に、歪み  $E_A(k)$ 、 $E_B(k)$  を比較し、これらの比較結果をフレーム内の全サブフレームに亘って判断処理することにより符号帳を切換選択するものである。

【0164】この第2の方法の実現例を図11に示す。この図11では、各サブフレーム毎に比較を行う比較器 516からの出力を判断ロジック 519に送って、例えば多数決などにより判断処理し、1ビットの符号帳切換選択フラグを端子 503より得るようにしている。

【0165】なお、この選択フラグが、前述した S/M (音声/楽音) モードデータとして伝送されるものである。

【0166】このようにして、複数の性質の異なる信号を、1つの量子化装置により効率よく量子化できる。

【0167】次に、図1のFFT回路161、周波数シフト回路162、逆FFT回路163による周波数変換処理について説明する。

【0168】この周波数変換処理は、入力信号の内の少なくとも1つの帯域を取り出す帯域抽出工程と、抽出された少なくとも1つの帯域の信号を周波数軸上の信号に変換する直交変換工程と、直交変換された信号を周波数軸上で(他の位置に、他の帯域に)シフトさせるシフト工程と、周波数軸上でシフトされた信号を逆直交変換して時間軸上の信号に変換する逆直交変換工程とを有している。

【0169】図12は、上記周波数変換のための構成をより詳しく示した図であり、図1と対応する部分には同じ番号を付している。この図12において、入力端子101には、例えば16kHzサンプリングで0~8kHzの成分を持つ広帯域音声信号が供給されている。この入力端子101からの広帯域音声信号の内、例えば0~3.8kHzを低域側信号としてLPF(ローパスフィルタ)102により分離し、また元の広帯域信号からこの低域側信号を減算器151で差し引いた成分を高域側信号として分離する。これらの低域側信号と高域側信号とを独立に処理するようにしている。

【0170】ここで得た高域側信号は、LPF102を介してもわずかに残っている3.5kHzから8kHzまでの4.5kHzの周波数幅を持つが、ダウンサンプリングして信号処理を行うため、4kHz幅まで狭めなくてはならない。この例では、7.5kHz~8kHzの0.5kHz分をBPF(バンドパスフィルタ)107あるいはLPFによりカットしている。

【0171】次に、低域側への周波数変換として例えば

FFT(高速フーリエ変換)を行うが、これに先立って、フレーム分割回路108により、サンプル数を2のべき乗、例えば図13の(A)に示すように512サンプル毎に区切っている。ただし、後での信号処理を行い易くするため、80サンプル毎に前進させている。

【0172】次に、ハミング窓かけ回路109により、長さ320サンプルのハミング窓をかける。このサンプル数の320は、上記フレーム分割の際に80サンプルずつ前進させており、後の重畳加算によるフレーム合成時に、図13の(B)に示すように4つの波形を重ねて加算できるようにするため、80の4倍としているものである。

【0173】次に、この長さ512サンプルのデータに対して、FFT回路161によりFFT処理を行い、周波数軸上のデータに変換する。

【0174】次に、周波数シフト回路162により、周波数軸上でデータを他の位置あるいは他の帯域にシフトあるいは移動させる。この周波数軸上でのシフトによってサンプリング周波数を低下させる原理は、図14に示すように、(A)の斜線部の高域側信号を(B)のように低域側に移動し、これを(C)に示すようにダウンサンプリングするものである。図14の(A)から(B)への周波数軸上での移動の際に、 $f_s/2$ を中心に折り返される成分については、互いに逆の移動方向となっている。これによって、サブバンドの帯域が $f_s/2n$ 以内であれば、サンプリング周波数を $f_s/n$ に下げることができる。

【0175】この周波数シフト回路162では、図15に示すように、周波数軸上のデータの高域側に相当する斜線部のデータを、低域側に相当する周波数軸上の位置あるいは帯域にシフトあるいは移動させる処理を行えばよい。具体的に、512サンプルの時間軸上のデータをFFT処理して得られる周波数軸上の512個のデータに対して、113番目から239番目までの127個のデータを、1~127番目の位置あるいは帯域に移動させ、また273番目から399番目までの127個のデータを、395~511番目の位置あるいは帯域に移動させる。このとき、周波数軸上の112番目のデータを0番目の位置に移動させないことが重要である。これは、周波数領域の信号では、0番目は直流成分であり、位相成分が無いため、この位置のデータは実数でなくてはならず、一般に複素数である周波数成分は、ここには入れられないからである。また、 $f_s/2$ を表す256番目のデータ(一般には $N/2$ 番目のデータ)も無効であり、ここは利用しない、すなわち、この場合の0~4kHzの範囲は、厳密には $0 < f < 4$ kHzと表される範囲のことである。

【0176】次に、逆FFT回路163により逆FFT処理して、周波数軸上のデータを時間軸上の信号に戻す。この場合512サンプル毎に時間軸上の信号が得ら

れる。この 512 サンプル毎の時間軸信号を、重畳加算回路 166 により図 j の (B) に示すように 80 サンプルずつオーバーラップさせ、重なっている部分を足し合わせる。

【0177】この重畳加算回路 166 から得られた信号は、16 kHz サンプリグで 0 ~ 4 kHz に制限されているので、ダウンサンプリグ回路 164 によりダウンサンプリグ処理する。これにより周波数シフトされた 8 kHz サンプリグで 0 ~ 4 kHz の信号を得ることができ、この信号が出力端子 169 を介して取り出されて、図 1 の LPC 分析・量子化部 130 や LPC 逆フィルタ 171 に送られる。

【0178】次に、デコード側での復元処理は、図 16 に示す構成により実現できる。

【0179】この図 16 の構成は、上記図 2 のアップサンプリグ回路 233 以降の構成に相当しており、対応する部分に同じ指示符号を付している。ただし、図 2 においては、FFT 処理の前にアップサンプリグ処理を行っているが、図 16 の例では、FFT 処理の後にアップサンプリグ処理を行っている。

【0180】この図 16 において、端子 241 には、図 2 の高域側 LPC 合成フィルタ 232 からの出力信号のような、8 kHz サンプリグで 0 ~ 4 kHz にシフトされている高域側の信号が入力される。

【0181】この信号は、フレーム分割回路 242 により、フレーム長が 256 サンプルで前進分が 80 サンプルの信号に区切られる。これは、エンコード側のフレーム分割と同様な理由からであるが、サンプリグ周波数が  $1/2$  となっているので、サンプル数も  $1/2$  となっている。また、ハミング窓かけ回路 243 により、フレーム分割回路 242 からの信号に長さ 160 サンプルのハミング窓がかけられることも、エンコード側と同様（ただしサンプル数は  $1/2$ ）である。

【0182】次に、FFT 回路 234 により長さ 256 サンプルで FFT 処理が施され、時間軸上の信号が周波数軸上の信号に変換される。次のアップサンプリグ回路 244 においては、図 15 の (B) に示すようなゼロ埋め処理を施すことにより、実質的にフレーム長が 216 サンプルから 512 サンプルになる。これは、図 14 の (C) から (B) への変換に相当する。

【0183】次に、周波数シフト回路 235 により、周波数軸上でデータを他の位置あるいは他の帯域にシフトあるいは移動させることにより、+3.5 kHz の周波数シフトを行う。これは、図 14 の (B) から (A) への変換に相当する。

【0184】このようにして得られた周波数軸上の信号を、逆 FFT 回路 236 により逆 FFT 処理することにより、時間軸上の信号に戻す。この逆 FFT 回路 236 からの信号は、16 kHz サンプリグで 3.5 kHz ~ 7.5 kHz となっている。

【0185】次の重畳加算回路 237 では、長さ 512 サンプルのフレーム毎に、80 サンプルずつオーバーラップさせて足し合わせ、連続する時間軸信号に戻す。このようにして得られた高域側信号は、加算器 228 で低域側信号と加算され、出力端子 229 より取り出される。

【0186】なお、このような周波数変換においては、具体的な数値は上記の例に限定されず、シフトを行うバンド数も 1 つに限定されない。

10 【0187】例えば、図 17 に示すように、16 kHz サンプリグで狭帯域信号を 300 Hz ~ 3.4 kHz、広帯域信号を 0 ~ 7 kHz とする場合に、狭帯域に含まれない低域 0 ~ 300 Hz と、高域側の 3.4 kHz ~ 7 kHz との内、高域側を 300 Hz ~ 3.9 kHz に移動して低域側と接触するように集めれば、0 ~ 3.9 kHz の信号となり、これも上述と同様にサンプリグ周波数  $f_s$  を  $1/2$  の 8 kHz とすることができる。

20 【0188】これを一般化すれば、広帯域信号と、広帯域信号の内部に収まる狭帯域信号とを多重化する場合、広帯域信号から狭帯域信号を減算した残りの内の高域側成分を低域側にシフトして、サンプリグレートを下げるわけである。

【0189】このように、任意の周波数から任意の周波数のサブバンドを作ることができ、その周波数幅の総和の 2 倍のサンプリグ周波数で処理可能であり、アプリケーションに柔軟に対応できる。

30 【0190】また、低ビットレートで量子化誤差が大きい場合、一般に QMF を利用すると分割周波数付近に発生したはずの折り返しノイズも、上記周波数変換方法によれば回避できる、という利点もある。

【0191】なお、本発明は上記実施の形態のみに限定されるものではなく、例えば上記図 1 の音声符号化側（エンコード側）の構成や、図 2 の音声復号化側（デコード側）の構成については、各部をハードウェア的に記載しているが、いわゆる DSP（デジタル信号プロセッサ）等を用いてソフトウェアプログラムにより実現することも可能である。また、上記ベクトル量子化の代わりに、複数フレームのデータをまとめてマトリクス量子化を施してもよい。さらに、本発明が適用される音声符号化方法や復号化方法は、上記符号化復号化方法に限定されるものではなく、種々の音声符号化復号化方法に適用でき、用途としても、伝送や記録再生に限定されず、ピッチ変換やスピード変換、規則音声合成、あるいは雑音抑圧のような種々の用途に応用できることは勿論である。

【0192】

40 【発明の効果】以上の説明から明らかなように、本発明に係る音声符号化方法によれば、入力信号を帯域分割し、分割された少なくとも 1 つの高域側の信号を低域側に周波数変換し、低域側に変換された信号のサンプリン

グレートが低下させ、サンプリングレートが低下させられた信号を予測符号化することにより、符号化効率が向上し、低ビットレートで高品質な符号化が可能となる。また、低域側と高域側とに分離して符号化することにより、広帯域の信号再生が行える。

#### 【図面の簡単な説明】

【図 1】本発明に係る音声符号化方法の実施の形態が適用される音声信号符号化装置の基本構成を示すブロック図である。

【図 2】音声信号復号化装置の基本構成を示すブロック図である。

【図 3】他の音声信号符号化装置の構成を示すブロック図である。

【図 4】伝送される符号化データのビットストリームのスケラビリティを説明するための図である。

【図 5】本発明が適用可能な符号化側のシステム全体を概略的に示すブロック図である。

【図 6】符号化、復号化の主要動作の周期及び位相関係を説明するための図である。

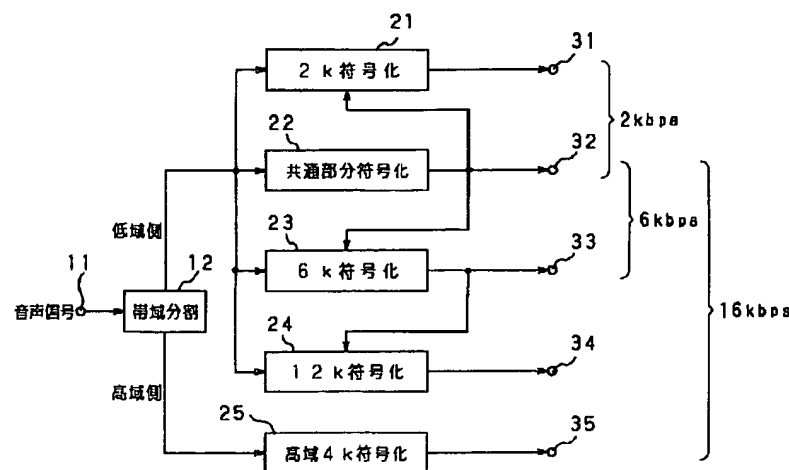
【図 7】MDCT（モディファイド離散コサイン変換）係数のベクトル量子化の構成例を示す図である。

【図 8】ポストフィルタ出力にかけられる窓関数の例を示す図である。

【図 9】2種類のコードブックを有するベクトル量子化装置の例を示す図である。

【図 10】2種類のコードブックを有するベクトル量子化装置の具体例を示す図である。

【図 5】



【図 11】2種類のコードブックを有するベクトル量子化装置の他の具体例を示す図である。

【図 12】周波数変換のエンコーダ側の構成を示すブロック図である。

【図 13】フレーム分割及び重畳加算処理を説明するための図である。

【図 14】周波数軸上での周波数シフトの他の例を示す図である。

【図 15】周波数軸上のデータのシフト処理を示す図である。

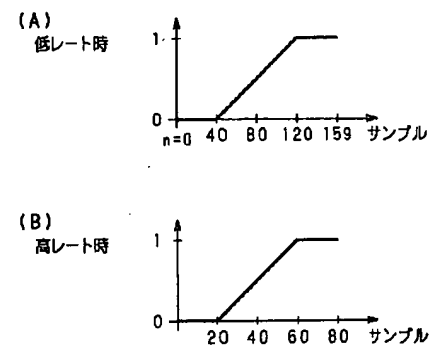
【図 16】周波数変換のデコーダ側の構成を示すブロック図である。

【図 17】周波数軸上での周波数シフトの他の例を示す図である。

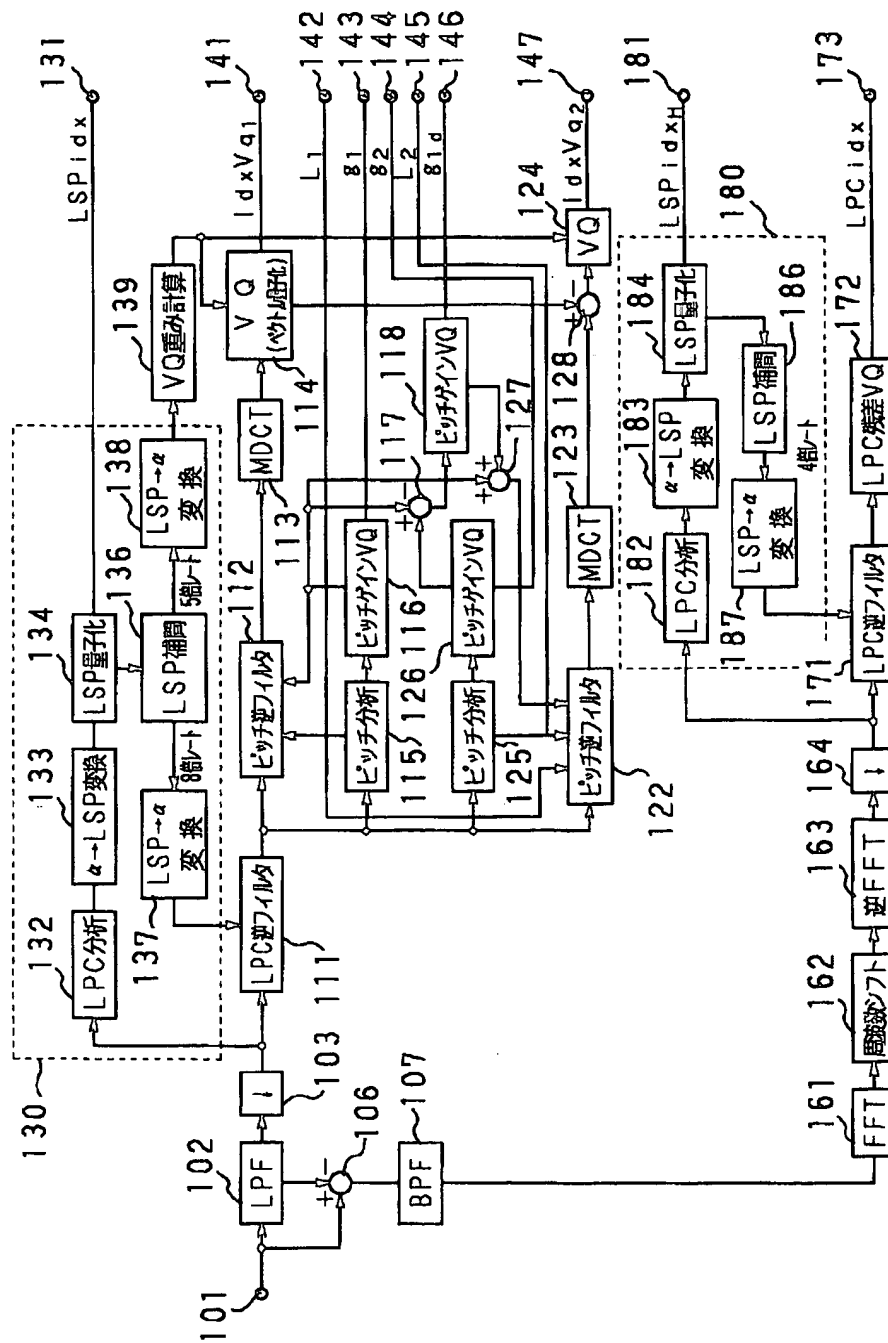
#### 【符号の説明】

- 111、171 LPC逆フィルタ
- 112、122 ピッチ逆フィルタ
- 113、123 MDCT（モディファイド離散コサイン変換）回路
- 114、124 VQ（ベクトル量子化）回路
- 115、125 ピッチ分析回路
- 116、118、126 ピッチゲインVQ回路
- 130、180 LPC分析・量子化部
- 161、234 FFT（高速フーリエ変換）回路
- 162、235 周波数シフト回路
- 163、236 逆FFT回路

【図 8】

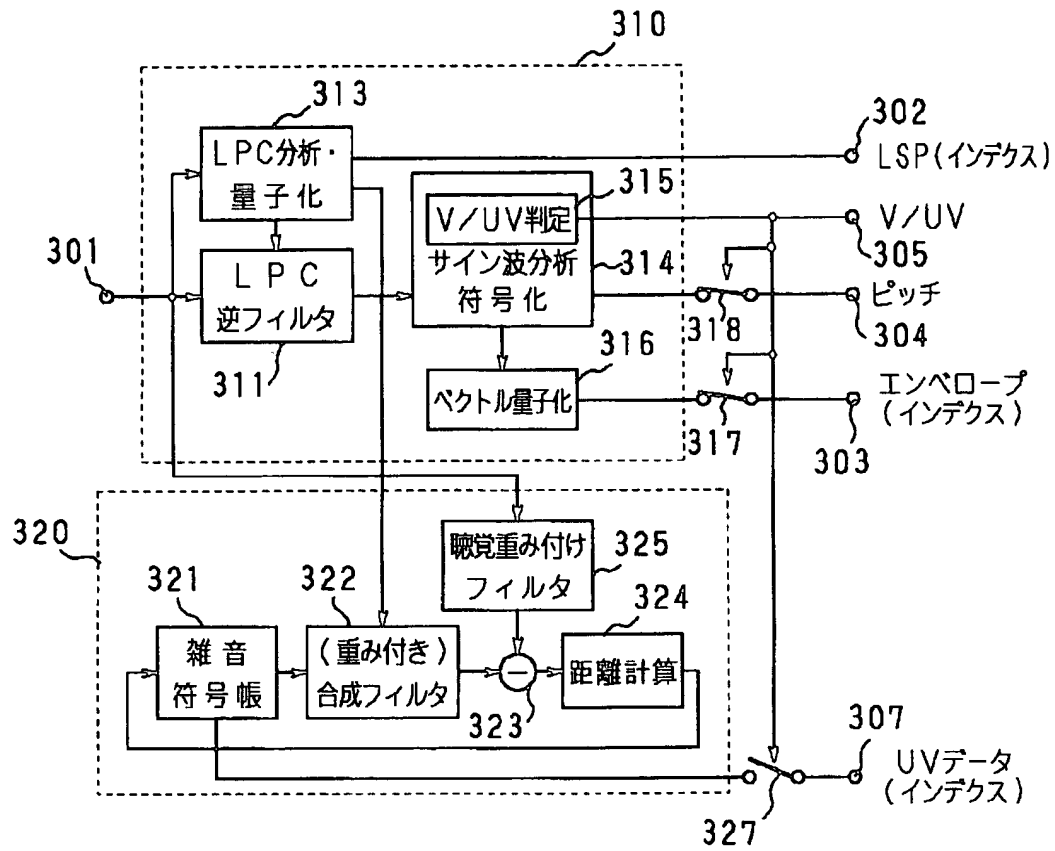


【图 1】

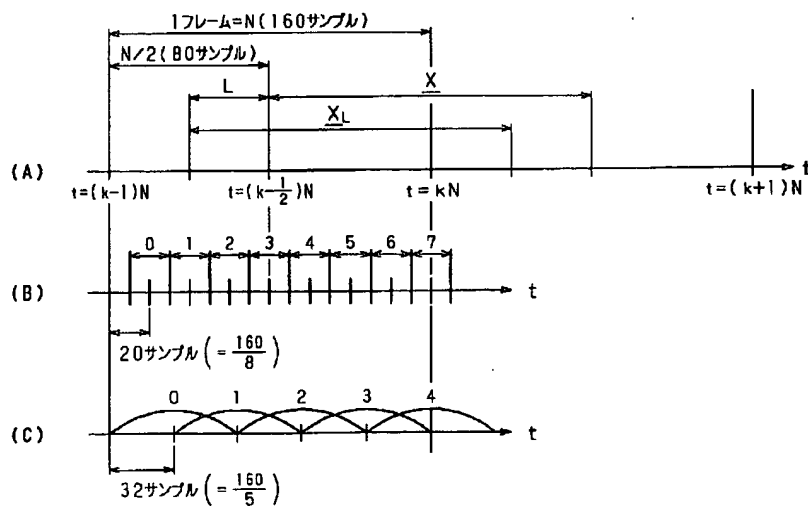




【図3】

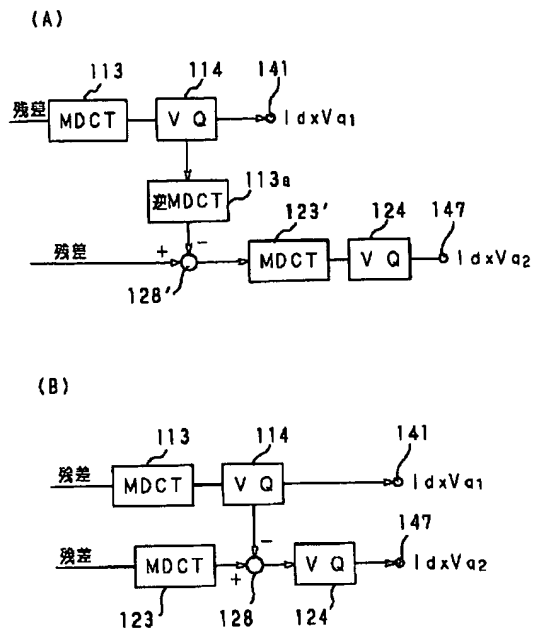


【図6】

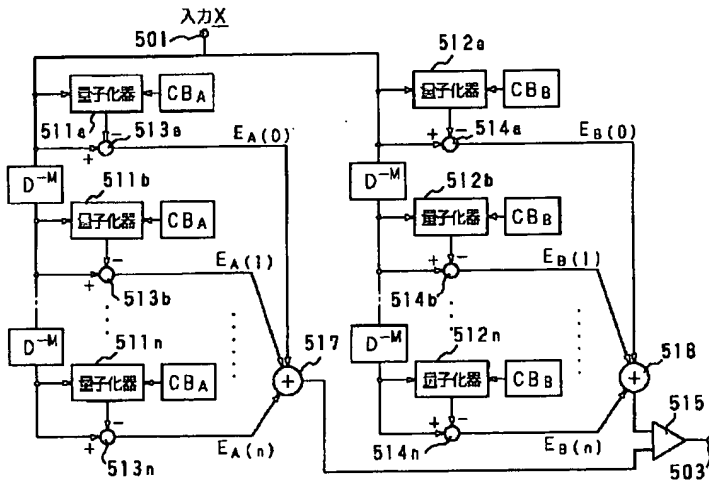




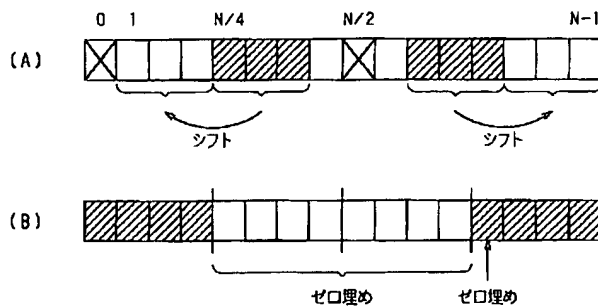
【図 7】



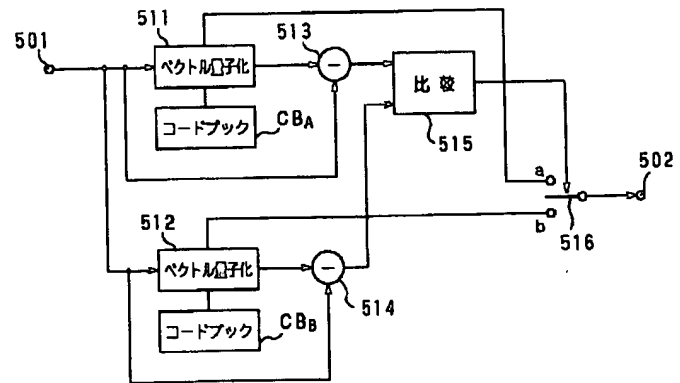
【図 10】



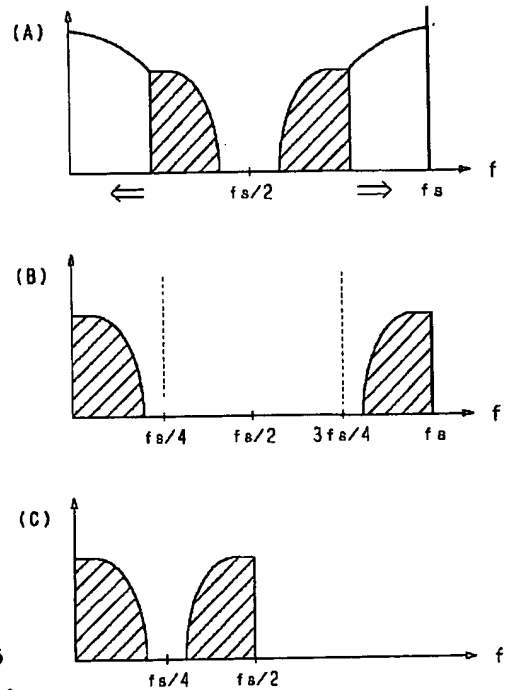
【図 15】



【図 9】



【図 14】







フロントページの続き

(72) 発明者 飯島 和幸

東京都品川区北品川 6 丁目 7 番 35 号 ソニ  
ー株式会社内